

PATENT ABSTRACTS OF JAPAN

(11)Publication number : 07-184160

(43)Date of publication of application : 21.07.1995

(51)Int.Cl.

H04N 5/91

H04N 5/781

(21)Application number : 05-328033

(71)Applicant : CANON INC

(22)Date of filing : 24.12.1993

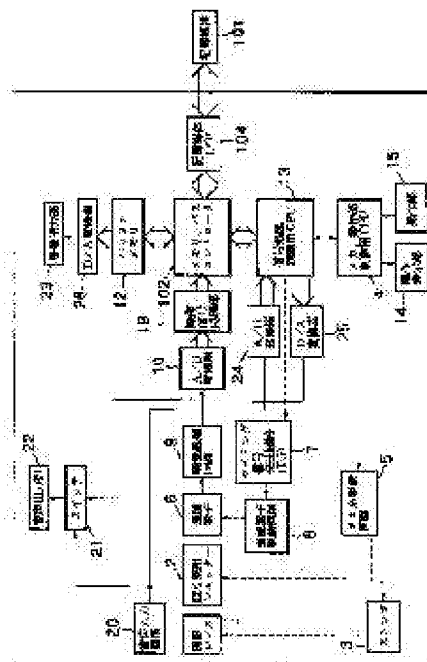
(72)Inventor : YAMAGAMI MIGAKU

(54) DEVICE FOR PROCESSING PICTURE DATA AND AUDIO DATA

(57)Abstract:

PURPOSE: To easily edit and record the expression in relation among a picture file, an audio file and a text file by designating audio data stored in a recording medium, retrieving the picture data relating to the audio data as file information and reading the data.

CONSTITUTION: When a signal processing CPU 13 detects a recording mode transition instruction of an operation section 15 by the user, the CPU 13 executes the following processing and displays a recorded picture to a video output section 23. A mechanism operation section CPU 4 and a drive circuit 5 control a lens system. The circuit 5 drives a shutter 2 based on a control variable from the CPU 4. Furthermore, a stroboscopic lamp 3 is lighted for the image pickup. A light of an object is made incident on an image pickup element 6 via a processing circuit 9 and an A/D converter section 10. A video signal of the element 6 is converted into a standard component video signal by a processing section 19 via a processing circuit 9 and an A/D converter section 10 and the video signal is inputted to a controller 102. The controller 102 displays a moving picture on a video output section 23 via a buffer memory 12 and a D/A converter 26 under the control of the control CPU 13 and records the picture on a recording medium 101 via an I/F 104.



(11)特許出願公開番号

特開平7-184160

(43)公開日 平成7年(1995)7月21日

(51) Int.Cl.⁶

識別記号

室内整理番号

FI

技術表示箇所

H04N 5/91

5/781

C

H04N 5/91

R

審査請求 未請求 請求項の数4 O.L (全 13 頁)

(21)出願番号 特願平5-328033

(22) 出願日 平成5年(1993)12月24日

(71)出願人 000001007

キヤノン株式会社

東京都大田区下丸子3丁目30番2号

(72) 尧明者 山上 琢

東京都大田区下丸子3丁目30番2号 キヤ
ノン株式会社内

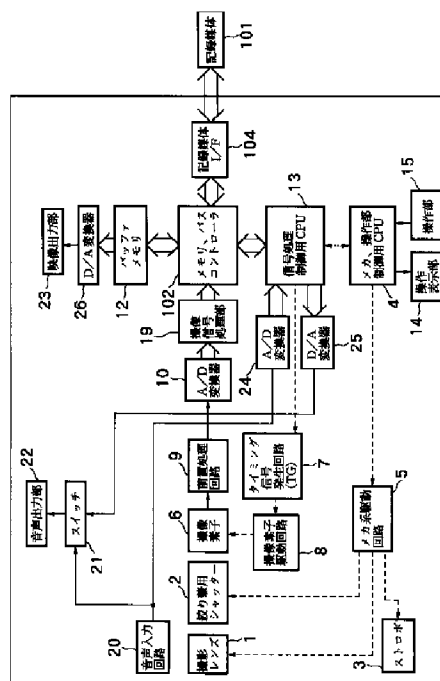
(74)代理人 弁理士 大塚 康德 (外1名)

(54) 【発明の名称】 画像データ及び音声データを処理する装置

(57) 【要約】

【目的】少なくとも1つ以上の画像、音声、テキストの各ファイル間のリンクが可能となり、画像、音声、テキストの各ファイル間での相互に関連性のある表現を容易に編集、記録できるデジタル電子カメラを提供することである。

【構成】画像データと音声データとを関連づけて記録するための記録媒体１０１と、記録媒体１０１に格納された音声データを指定する指定手段２０３と、指定された音声データに関連する画像データをファイル情報として検索する制御手段１３と、検索された画像データを読み出し、再生する読出手段１０２とを具備することを特徴とする画像データ及び音声データを処理する装置。



【特許請求の範囲】

【請求項 1】 画像データと音声データとを関連づけて記録するための記録媒体と、前記記録媒体に格納された音声データを指定する指定手段と、前記指定された音声データに関連する画像データをファイル情報として検索する制御手段と、前記検索された画像データを読み出し、再生する読出手段とを具備することを特徴とする画像データ及び音声データを処理する装置。

【請求項 2】 前記画像データから文字認識を行なう文字認識手段を更に具備し、該文字認識手段による認識結果をテキストファイルとして前記画像及び音声の各データとに関連づけて記録し、かつ検索できるように登録処理を行うことを特徴とする請求項 1 に記載の画像データ及び音声データを処理する装置。

【請求項 3】 前記音声データを文字として認識する音声認識処理手段を更に具備し、該音声認識処理手段による認識結果をテキストファイルとして前記画像及び音声の各データとに関連づけて記録し、かつ検索できるように登録処理を行うことを特徴とする請求項 1 に記載の画像データ及び音声データを処理する装置。

【請求項 4】 前記記録媒体に格納された前記各データをファイル情報として表示する表示手段を更に具備し、前記各ファイル間での相互関連情報の編集、及び該関連情報に基づく前記各データの検索結果を表示することを特徴とする請求項 1 に記載の画像データ及び音声データを処理する装置。

【発明の詳細な説明】

【0001】

【産業上の利用分野】 本発明は画像及び音声の処理して、記録再生を行なう装置に関するものである。

【0002】

【従来の技術】 従来、画像及び音声データを処理し、記録する規格としてスチルビデオフォーマット方式が知られている。この従来のスチルビデオフォーマットは、画像データと音声データは夫々別のトラックに FM 記録される。また、音声トラックは関連する画像トラックのトラック番号を記録するためのフィールドをコントロールコードに設けられており、特定の画像についての注釈（識別情報）を記録できるようにしたものである。

【0003】

【発明が解決しようとする問題点】 しかしながら、上記のような従来のデジタル電子カメラにおいては、以下のような問題がある。

①即ち、スチルビデオフォーマット方式では、音声トラックからの参照したい画像トラックの指定はできるが、画像トラックからは関連する音声トラックを指定できない。従って、音声データと画像データとの関連を確定するためには記録媒体であるスチルビデオフロッピー上の

全てのトラックを検索して全ての音声トラックを再生し、参照したい画像トラックを見つけ出さなければならず必然的に多大な時間を必要とするため無駄が多く現実的ではないという欠点があった。

【0004】 ②また、音声トラックからは一つの画像しか参照できないので、複数の画像である場合に一つの音声に対応させることができなかった。従って、複数の画像に対して、共通の概念を音声で注釈しようとしても一度にできないという欠点があった。

③更に、音声データを記録するためには多大な記録容量を必要とするので、特に単純な注釈を音声トラックにて行うことはコスト的にみて不適切であるという問題があった。

【0005】 尚、上述した従来のデジタル電子カメラに関する問題点は、一般的な音声及び画像データを処理する装置に係わる問題点でもある。従って、本発明は上述した課題に鑑みてなされたものであり、その目的とするところは、少なくとも 1 つ以上の画像、音声、テキストの各ファイル間のリンクを可能とし、それぞれのファイルを双方から高速に検索、再生できるように構成して、画像、音声、テキストファイル間での相互に関連性のある表現を容易に編集、記録できる音声及び画像データを処理する装置を提供することである。

【0006】 また、音声あるいは画像をテキスト形式に変換できるため少ない容量で画像に注釈（識別情報）を付与することができる音声及び画像データを処理する装置を提供することである。また、注釈として付与されたテキストに基づいて検索を行ったり、データベースを作成するときのキーワードとして使用することも可能となる音声及び画像データを処理する装置を提供することである。

【0007】 更に、他のデータベースに格納された情報との間で自動的にデータをリンクするなどという処理も可能となる音声及び画像データを処理する装置を提供することである。

【0008】

【課題を解決するための手段】 上述の課題を解決し、目的を達成するために、本発明の画像データ及び音声データを処理する装置は、画像データと音声データとを関連づけて記録するための記録媒体と、前記記録媒体に格納された音声データを指定する指定手段と、前記指定された音声データに関連する画像データをファイル情報として検索する制御手段と、前記検索された画像データを読み出し、再生する読出手段とを具備することを特徴としている。

【0009】 また、好ましくは前記画像データから文字認識を行なう文字認識手段を更に具備し、該文字認識手段による認識結果をテキストファイルとして前記画像及び音声の各データとに関連づけて記録し、かつ検索できるように登録処理を行うことを特徴としている。また、

好ましくは前記音声データを文字として認識する音声認識処理手段を更に具備し、該音声認識処理手段による認識結果をテキストファイルとして前記画像及び音声の各データとに関連づけて記録し、かつ検索できるように登録処理を行うことを特徴としている。

【0010】また、好ましくは前記記録媒体に格納された前記各データをファイル情報として表示する表示手段を更に具備し、前記各ファイル間での相互関連情報の編集、及び該関連情報に基づく前記各データの検索結果を表示することを特徴としている。

【0011】

【作用】以上のように、この発明に係わる画像データ及び音声データを処理する装置は構成されているので、少なくとも1つ以上の画像、音声、テキストの各ファイル間のリンクが可能となり、それぞれのファイルを双方から高速に検索、再生できるので、画像、音声、テキストファイル間での相互に関連性のある表現を容易に編集、記録できる。

【0012】また、音声あるいは画像をテキスト形式に変換できるため少ない容量で画像に注釈（識別情報）を付与することができる。また、注釈として付与されたテキストに基づいて検索を行ったり、データベースを作成するときのキーワードとして使用することも可能となる。更に、他のデータベースに格納された情報との間で自動的にデータをリンクするなどという処理も可能となる。

【0013】

【実施例】以下、本発明の好適な実施例について、添付図面を参照して詳細に説明する。図1は、本実施例のデジタル電子カメラ本体のシステム構成を示すブロック図である。図1において、記録媒体101は、例えばPCMCIA規格に準拠したメモリカードやハードディスクなどである。音声入力回路20及び音声出力部22は、例えばオーディオジャックあるいはスピーカである。A/D変換器24は、音声信号をアナログ信号からデジタル信号に変換する装置で、D/A変換器25は、信号処理制御用CPU13から送られるデジタルの音声信号をアナログに変換する装置である。スイッチ21は音声出力部22への音声信号の送信を選択する選択回路である。メモリバスコントローラ102は、撮像信号処理部19、信号処理制御用DSP13、画像表示用バッファメモリ12、記録媒体I/F回路104の間での画像データ、音声データなどの転送を行う。

【0014】D/A変換機26は画像表示用バッファメモリ12からのデジタル画像データをアナログ映像信号に変換する装置であり、映像出力部23はアナログに変換された映像信号を映像として表示する映像表示装置である。1は撮影レンズであり、2は絞り機能とシャッター機能を兼ねる絞り兼用シャッター、3はストロボ、メカ及び操作部制御用CPU4は機械的な操作を行なう各

操作部を制御し、駆動回路5はメカ系各部の駆動させる回路である。

【0015】撮像素子6は被写体からの反射光を電気信号に変換するCCDで、タイミング信号発生回路7は撮像素子6を動作させるために必要なタイミング信号を発生する（以下、「TG」と称する）。撮像素子駆動回路8はタイミング信号発生回路7からの信号を撮像信号の駆動可能なレベルに増幅する駆動回路で、前置処理回路9は撮像素子6で発生する出力ノイズ除去のためのCDS回路及びA/D変換前に行う非線形増幅回路を備え

る。A/D変換器10は前置処理後のデータをデジタルに変換する。信号処理制御用CPU13は信号処理部を制御する信号処理制御用DSP（DIGITAL SIGNAL PROCESSOR）であり、操作表示部14は操作補助のための表示やカメラの状態を表わす表示部で、操作部15はカメラを外部から制御するためのキーボード等の入力装置である。記録媒体I/F104は本実施例に基づくデジタル電子カメラと記録媒体101とを接続するための記録媒体I/F用回路である。そして、本実施例のカメラの記録媒体への記録ファイルフォーマットは、例えばMS-DOSのようなファイルフォーマットを使用することができる。

【0016】画像表示用バッファメモリ12は信号処理制御用DSP13から画素単位でアクセスすることが可能であり、また撮影する画像を表示しながら任意の操作手段パネルを描画することができる。操作部15にトラックボールなどを実装すれば、近年におけるパーソナルコンピュータにおけるポインティングデバイスを用いたGUI（Graphical User Interface）と同様のユーザーI/Fを実現できるようになっている。つまり信号処理制御用DSP13が画像表示用バッファに様々なコントロールパネルを描画して、ユーザーがそのコントロールパネルに対してポインティングデバイス（この場合トラックボール）によって操作することでカメラのあらゆる操作が可能となる。また信号処理制御用DSP13によって、画像やテキストや様々な図形を画像表示部に描画することでユーザーに対して撮影された画像、カメラの現在の状態、記録媒体上のファイル管理情報など多種多様な情報の提示が行われる。本実施例では以下操作部15と上記GUIによるユーザーI/F操作を操作部15による操作と称することとする。またポインティングデバイスによるクリックあるいはダブルクリック操作などを起動操作と称することとする。

【0017】＜画像記録モード時の記録画像のモニタ＞信号処理制御用DSP13がユーザーの操作部15の記録モード移行命令を検知すると、信号処理制御用DSP13は以下の処理を実行し記録する画像を映像出力部23に表示する。撮影者の意図に応じてレンズ系の制御がメカ操作部制御用CPU4及びメカ系駆動回路5により行われる。この際、撮影条件などが操作部15に表示さ

れ、撮影者にカメラの状況を伝える。さらに不図示の測光回路により被写体の明るさを測定し、絞り兼用シャッター2の絞り度を表す値やシャッタースピードのデータをメカ操作部制御用CPUにて導出する。メカ操作部制御用CPU4で導出された制御値に基づき、メカ系駆動回路5により絞り兼用シャッター2を駆動する。また測光回路（不図示）の出力によってはストロボ3を発光させて撮影することになる。このようにして露光されて、被写体の反射光が撮影レンズ1及び絞り兼用シャッター2を介して撮像素子6に入射される。この際、絞り兼用シャッター2は撮像素子6への入射光量を制限するとともに、撮像素子としてインターレース読み出し型CCDを用いた場合、画像の転送中に入射光が信号電荷に悪影響を与えないようにするために設けられている。撮像素子6は、TG7からの出力を撮像素子駆動回路8によって増幅させた駆動信号により動作させる。尚、TG7は信号処理制御用DSP13によりその動作を制御されている。以上のようにして駆動させた撮像素子6の出力は前置処理回路9に送られる。前置処理回路9では撮像素子からの出力に発生する低域ノイズを除去するCDS処理と共に、A/D変換器10のDレンジ（デジタル化された信号データ）を有効に用いるために撮像出力を非線形化する処理を行っている。前置処理された撮像信号出力はA/D変換器10においてデジタル信号に変換され、撮像信号処理部19によって標準的なコンポーネントビデオ（例えば輝度信号と2色差信号やRGB信号）に変換されメモリコントローラ102に入力される。メモリコントローラ102では信号処理制御用DSP13の制御により、デジタル化された撮像信号をバッファメモリ12に連続的に転送していく。

【0018】図2は、画像記録モードにおける表示形式を示す図である。画像表示部23に出力される表示画像は例えば図2に示すように記録画像のモニタ領域部分202と、カメラの状態や撮影条件などユーザーが設定できるパラメータのコントロールパネル部分203を分割されて表示される。画像表示用バッファメモリ12のデジタル化されたデータはA/D変換器26によってアナログ映像信号に変換され映像出力部23に出力され、映像出力部23で表示される。ユーザーは、記録する画像を動画像としてモニタ領域部分202において確認できる。

【0019】以上の処理によって撮像素子6によって捕らえられた動画像が映像出力部23に表示されるのである。

<画像の記録>撮影者が操作部15を制御することによりカメラに撮影を指示すると、信号処理制御用DSP13は動画表示を停止して表示用バッファメモリに保持された画像データにメモリコントローラを介してアクセスしてデジタルの圧縮処理を行った後、記録媒体I/F104を介して記録媒体101に記録する。

【0020】この転送期間及び記録終了後一定期間中メモリバスコントローラ102は画像表示用バッファメモリ12への書き込みを停止すれば記録する画像が静止されて映像出力部23に表示される。従ってユーザーがたった今記録した静止画像を映像出力部23で確認できる。また撮影後、より速く記録された画像の中身を確認するために、信号処理制御用DSP13は圧縮した画像の間引き画像を圧縮ファイルに付加することができる。例えば、元の画像の縦横8分の1程度に間引いた画像を付加してもファイル容量はそれほど増加しないのである。尚、この画像を索引画像と呼ぶ。

【0021】<音声記録時の音声のモニタ及び記録>記録モード時には信号処理制御用DSP13はスイッチ回路21で音声入力回路20の出力を音声出力部22の入力に接続して記録する音声をモニタすることができる。信号処理制御用DSP13がユーザーの操作部15による音声記録命令を検知すると、信号処理制御用DSP13はA/D変換器24でデジタルデータに変換したデータを受け取り、メモリバスコントローラ102を介して記録媒体I/F104に転送する。

【0022】信号処理制御用DSP13がユーザーの操作部15による音声記録命令の解除を検知するか、あるいは信号処理制御用DSP13が一定時間経過したことで音声記録の解除とみなした時点で信号処理制御用DSP13は音声の記録を終了する。

<画像と音声の同時記録>画像と音声を同時に記録する場合、前述の<画像の記録>で説明したように画像を圧縮して記録媒体I/F104に転送している間は、信号処理制御用DSP13はA/D変換器24から受け取った音声のデータを信号処理制御用DSP13の内部バッファに一時的に保存しておき、画像の転送が終了した時点で音声データを記録媒体I/F104に転送する。NTSC方式の1フィールド期間16.7ミリ秒のうち、垂直同期期間とその前後には1.4ミリ秒程度の垂直ブランキング期間といわれる映像信号が存在しない期間がある。画像データを通常のビデオレートで転送する場合、この垂直ブランキング期間を除いた時間（約15ミリ秒）の間に撮像信号処理されていない画像データを転送する。これを実現するために約10MByte/secのスピードでデータを転送する。

【0023】音声データを1サンプル8ビットで22kHzでサンプリングするとき16.7ミリ秒分のデータ容量は約370Byteである。このデータを残りの1.4ミリ秒で転送する場合、約260KByte/secの転送スピードが必要である。これらの転送スピードは、例えば記録媒体としてPCMCIA規格などに準拠したメモ리카ードならば十分満足できる速度である。

【0024】以上説明したように、1画面の走査期間のうち垂直ブランキング期間中に音声データを転送し、映像期間に画像転送を行い時分割多重することで本実施例

のカメラは画像と音声を同時にモニタしながら同時に記録できる。このときは音声データは画像に関連した情報とみなせるため、画像及び音声を後述する<画像ファイルへの音声及びテキストのリンク>の説明に基づいて、両者が相互に参照し合うようにファイル管理する。

【0025】また音声データを映像信号の水平ブランキング期間内に転送しても良い。例えば、44.1kHz、ステレオ、8ビットの条件で音声をサンプリングするとき、1回の水平ブランキング期間（1H期間）内に計6Byteもしくは4Byteの音声データを転送・記録する。この時、記録媒体としてHDDなどの記録機構を有する媒体を用いる際には1H期間毎に音声データがインターリーブされるように記録するのが良いが、半導体メモリの場合には画像データと音声データを1H毎にインターリーブさせる必要はなく、1フィールド期間（1V期間）の画像データの前に音声データが記録されるように1V毎のインターリーブ記録をしても良い。

【0026】<テキストデータの記録>本実施例のカメラでは5つの画像や音声を認識してテキストデータにすることで画像に対する注釈のための記録容量の拡大を大幅に削減する。

<①文字認識による記録>信号処理制御用DSP13がユーザーの操作表示部15による文字認識記録モード移行命令を感知するとカメラは<画像記録モード時の記録モニタ>と同じ処理を行い文字認識する画像を画像表示部に出力する。このとき画像表示部301に表示される画像は、図3のように認識しようとする画像302、文字認識された結果を表示する部分303、カメラの状態や撮影条件などをユーザーが設定できるパラメータのコントロールパネル304などを分割して表示する。

【0027】そしてユーザーが操作部15によって文字認識実行を指示している間は以下の処理を繰り返す。信号処理制御用DSP13は一旦画像表示用バッファメモリへの書き込みを停止してその画像に対して文字認識処理を行う。バッファメモリ上の画像に対する文字認識処理が終了すると信号処理制御用DSP13はその認識結果を図3に示す表示領域303に表示する。

【0028】ユーザーは満足な認識結果が得られたところで操作部15によって文字認識実行の指示を終了する。この時点でユーザーは操作部15によって認識されたテキストの確定を指示する。カメラは確定されたテキストをメモリコントローラ102、記録媒体I/F104を介して記録媒体101に記録する。テキストデータが確定されない場合は、再度認識画像をバッファメモリに書き込み直し上述の動作を繰り返す。

【0029】<②音声認識による記録>信号処理制御用DSP13がユーザーの操作部15による文字認識記録モード移行命令を感知するとカメラは<音声記録時の音声のモニタ及び記録>に説明したのと同じ処理を行い音声をモニタする。このとき画像表示部401の表示画像

は図4のように音声認識された結果を表示する部分402、カメラの状態や、認識条件等ユーザーが設定できるパラメータのコントロールパネル403などを分割して表示する。さらに好適な具体例として、図13のパネル130のように構成できる。

【0030】そしてユーザーが操作部15によって音声認識実行を指示している間信号処理制御用DSP13はA/D変換器24でデジタルデータに変換したデータを受け取り、音声認識処理を行い認識結果を図4の表示部402に表示する。ユーザーは音声の認識結果が得られたところで操作部15によって音声認識実行の指示を終了する。

【0031】認識結果に不満がある場合は再度操作部15によって音声認識実行を指示する。満足する認識結果が得られたらこの時点でユーザーは操作部15によって認識されたテキストの確定を指示する。カメラは確定されたテキストをメモリコントローラ102、記録媒体I/F104を介して記録媒体に記録する。

【0032】以上の説明により明らかなように本実施例のデジタル電子カメラは記録媒体に画像データ、音声データ、テキストデータの3種類のファイルを保持できる。

<ファイルの一覧表示>図5は、記録されたファイルの一覧を表示する表示形式を示す図である。3種のファイルをカメラの表示部に映像としてディレクトリ表示するときは信号処理制御用DSP13が記録媒体上に記録されているファイルを読み出し、ファイルデータの種類の応じて図5に示すように画像については画像ファイルを表現する索引画像501、音声、テキストは夫々502、504のようなアイコンとして表示する。これらの表示は記録時間順に表示したり、また画像だけ、音声だけ、テキストだけの表示を行うこともできる。またユーザーは索引画像、音声、テキストを表すアイコンをポインティングデバイスで選択したのち、消去ボタン509をクリックするとカメラは選択されたファイルの消去を行う。

【0033】索引画像501の下部にある音声、テキストの夫々のアイコン511、512はその画像ファイルにリンクされたテキストや音声が存在するかどうかを示すためのものである。テキスト及び音声は後述する<画像ファイルへの音声、テキストのリンク>に従ってリンクされる。例えば、511、512のようにアイコンをグレーレベルにすることでデータが存在することを表現する。これらのアイコンをポインティングデバイスを用いて選択して起動させ、後述する<音声の再生>又は<テキストの再生>が行われる。

【0034】図5において、画面の一部510は表示する索引画像やアイコンなどのオブジェクトが画面に入りきらないときにそれらを検索できるように上下にスクロールするためのスクロールバーである。これは近年のバ

ーソナルコンピュータやワークステーションで用いられるウィンドウ表示のためのツールと同様に機能する。

＜画像の拡大及び再生＞図7は、拡大された画像の表示形式を示す図である。図7において一つの画像について拡大再生したい場合、ユーザーは操作部のポインティングデバイスを用いて一つの索引画像を選択して拡大命令（例えばボタンのダブルクリック）を行う。信号処理制御用DSP13は上記操作を検知すると選択された画像ファイルから圧縮画像データを読み出し、伸長して画像表示用バッファメモリに転送し画像表示部に表示する。このときの表示は、例えば図7に示すように拡大された画像の表示部701と各制御ボタン702～704、更にリンクされた複数の音声、テキストのアイコンとして夫々705、706などが表示される。

【0035】ポインティングデバイスでボタン702を起動すると図7に表示された画像のパネルは閉じられて前述の図5の表示に戻る。

＜音声の再生＞次に、図5において、音声を再生したい場合、ユーザーは操作部のポインティングデバイスを用いて音声アイコンを選択して起動する。

【0036】信号処理制御用DSP13はスイッチ回路21を制御してD/A変換器25からの出力を音声出力部22の入力に接続する。信号処理制御用DSP13は記録媒体I/F104、メモリバスコントローラ102を制御して音声データを読み出しD/A変換器に記録時の標準化周期で出力し音声出力部22から音声が出力される。

【0037】＜テキストの再生＞更に、テキストを表示したい場合、ユーザは操作部のポインティングデバイスを用いてテキストアイコンを選択して起動する。図7は、テキストデータの表示形式を示す図である。信号処理制御用DSP13は記録媒体I/F104、メモリバスコントローラ102を制御してテキストデータを読み出しキャラクタビットパターン展開して例えば図6に示すようなテキストデータを画像表示用バッファメモリに転送し画像表示部に表示する。

【0038】図6ではテキストの表示部601と制御ボタン602などが表示される。ポインティングデバイスでボタン602を起動すると第6図のパネルは閉じられる。

＜画像ファイルへの音声及びテキストのリンク＞本実施例に基づくカメラは前述の＜単独の画像の拡大及び再生＞での説明に従って一つの画像を再生したとき、画像に対して音声、あるいは音声認識によるテキストの付加を行うためのコントロールボタンを図7に示すアイコン703、704を夫々備えている。ポインティングデバイスで図7中のアイコン703、704をクリックし起動すると信号処理制御用DSP13は夫々＜音声記録時の音声のモニタ及び記録＞、＜①音声認識による記録＞と同様に音声、テキストの記録を行う。この操作は画像デ

ータを拡大及び再生しモニタすることによって行うことができる。図13は、図7の画像表示パネルの上に音声認識の制御用パネル130を重ねて表示した様子を示す図である。例えば、図13のように図7の画像表示パネルの上に音声認識の制御用パネル130を重ねて表示することができる。画像データと音声データを同時記録する時、及び上記のように画像データを再生した結果をモニタしながら音声、テキストを記録する時は画像ファイル及び音声、テキストファイル双方がお互いを参照するデータをファイル内に格納することによって相互にリンクされたファイルとなる。図8は、ファイル内の画像ファイル及び音声ファイルとテキストファイル夫々のデータ構成を示している。

【0039】図8において、夫々のファイルはリンクされるファイル形式ごとにリンクされたファイル数（夫々音声ファイル数をn、テキストファイル数をm、画像ファイル数をk）及びファイル数の応じたファイルの識別子を格納する。このようにすれば音声ファイルだけでなく画像ファイルも音声、テキストファイルを参照しているので従来のように全てのファイルを検索しなくても一つの画像ファイルから関連される音声、テキストファイルを特定してそのファイルを再生あるいは表示することができる。

【0040】一つの画像ファイルにリンクされた音声、テキストが存在する場合、その画像を表示する場合夫々図5、図7に示すようにリンクされた音声ファイル、テキストファイルをアイコンで（例えば、画像データの下部に）表示することができる。ユーザーはこのアイコンをポインティングデバイスで選択、起動するとカメラは関連づけられた音声の再生、テキストの表示を行う。

【0041】テキストは図6のように表示される。この表示は単独、または図5、図7の表示の上に重ねて表示される。またユーザーは図7においてアイコン705、706をポインティングデバイスで選択した後、削除ボタン707をクリックするとカメラは関連づけられた音声、テキストファイルの削除を行うことが可能である。この操作手段は図8に示した画像ファイル、音声、テキストファイルのお互いのリンク情報を消去する。このときリンクの削除された音声やテキストファイルは単独で存在するようにしてもよいし、消去するようにしてもよい。

【0042】また、単独で表示させるには、図5においてアイコン502、504のファイル一覧表示のように夫々表現される。

＜複数の画像ファイルのグループ化、及びグループへの音声、テキストのリンク＞前述の＜ファイルの一覧表示＞において、ポインティングデバイスによって複数の画像を選択すると索引画像は選択されたことを強調するように例えば図9のように枠が太く表示される。更に、ポインティングデバイスで図9のグループ化ボタン901

をクリックすると画像ファイルのグループが作成される。この時グループ化されたことを強調するために索引画像の枠の色を他の画像と異なる色にすることもできる。

【0043】更に、ポインティングデバイスで図9に示す音声付加ボタン902あるいは音声テキスト付加ボタン903を起動すると信号処理制御用DSP13は夫々<音声記録時の音声のモニタ及び記録>、<②音声認識による記録>での説明と同様に音声、テキストの記録を行う。この時点で複数の音声、テキストがグループに付加されることとなる。この操作によってグループに特有な事柄についての注釈説明を与えることができる。

【0044】以上説明した操作により複数の画像と音声、テキストがリンクされる。この時グループ内の夫々のファイルについてそれ以外の全てのファイルに対する参照データを付加すると、のちにこのグループを特定するためには全てのファイルを検索しなければならず管理が困難になる。そこで本実施例ではグループの発生が明確になった時点でグループ情報を保持するためのグループファイルを作成する。このファイルの構成は、図10に示すようになる。即ち、グループに属する画像、音声、テキスト夫々のファイルの個数、及び夫々のファイルの識別子が格納される。またグループに属する画像、音声、テキストファイルは夫々のファイルが属するグループへの参照ができるように図11のように編成される。即ち、個々のファイルはグループを参照するだけでそのグループに属するファイルへの参照は行わない。

【0045】このようなグループのファイル構成により、グループ、画像、音声、テキスト、どのファイルからでもグループ情報すなわちグループに属するファイルを高速に同定できる。例えば一つの画像ファイルからそのグループに属する他のファイルを探索する場合、一旦所属するグループファイルを獲得してしまえばそれらのファイルの識別子を直接獲得できる。

【0046】以上説明した実施例から明らかなように画像ファイルは複数のグループに属することが可能である。上述の手順によってグループファイルが発生されるとカメラは<ファイルの一覧表>において、図5のようにグループファイルを表示するアイコン505を表示する。ポインティングデバイスでこのアイコンをダブルクリックすると図12のようにグループに属するファイルが表示される。図12において、音声アイコン、テキストアイコンをポインティングデバイスでクリックして起動すれば夫々前述した通りの<音声の再生>、<テキストの表示手段>が行われ、このグループに特有な事柄についての注釈、説明として表現することができる。

【0047】グループに属するファイルをグループから削除するためには図12において削除するファイルをポインティングデバイスで選択した後、ボタン121で削除を起動する。グループから削除された音声やテキスト

ファイルは単独で存在するようにする。単独で存在させるようにするときは図5においてアイコン501、502、504のように一覧表示に現れるようになる。

【0048】また消去ボタン122を設定してこのボタンを起動することで選択されたファイルを消去してもよい。グループから削除、あるいは消去される場合はもちろん、図10、図11に示されるグループファイル、画像、音声テキストデータファイルとの間のリンク情報は消去される。

【0049】またファイルを既存のグループに加えるためには図5において加えるファイル及び加える先のグループアイコンをポインティングデバイスで選択してグループ化を起動する。これらの操作においてグループ以外のファイルは複数選択することが可能である。記録媒体には単独で画像ファイル、音声ファイル、テキストファイルが存在する。これらのいくつかをグループ化することも可能である。<ファイルの一覧表>においてポインティングデバイスによって複数のファイルを選択する。さらにポインティングデバイスで第5図のグループボタン508を起動することで画像、音声、テキストが属するグループファイルを作成する。

【0050】(本実施例の効果)以上説明したように、本実施例においては、単一の画像又は複数の画像と、音声、テキスト間のリンクが可能になりそのリンクされた関係を高速に再現できるので、画像、音声、テキストの相互補間的な表現を容易に編集することができる。

【0051】また音声あるいは画像をテキストに変換するため少ない容量で画像に注釈を付けることができる。また注釈としてつけられたテキストをもとに検索を行ったり、データベースを形成するときのキーワードとすることも可能となる。さらには他のデータベースの情報と自動的にリンクするなどの処理も可能となる。

【0052】尚、本発明は、その趣旨を逸脱しない範囲で上記実施例を修正又は変形したものに適用可能である。例えば、複数の機器から構成されるシステムに適用しても1つの機器から成る装置に適用しても良い。また、システム或は装置にプログラムを供給することによって達成される場合にも適用できることはいうまでもない。

【0053】

【発明の効果】以上説明したように、本発明によれば、少なくとも1つ以上の画像、音声、テキストの各ファイル間のリンクが可能となり、それぞれのファイルを双方から高速に検索、再生できるので、画像、音声、テキストファイル間での相互に関連性のある表現を容易に編集、記録することができる効果がある。

【0054】また、音声あるいは画像をテキスト形式に変換できるため少ない容量で画像に注釈(識別情報)を付与することができる効果がある。また、注釈として付与されたテキストに基づいて検索を行ったり、データベ

ースを作成するときのキーワードとして使用することも可能となる効果がある。更に、他のデータベースに格納された情報との間で自動的にデータをリンクするなどという処理も可能となる効果がある。

【図面の簡単な説明】

【図 1】本発明の実施例のデジタル電子カメラの全体構成を示すブロック図である。

【図 2】画像記録モードにおける表示形式を示す図である。

【図 3】文字認識モードにおける表示形式を示す図である。

【図 4】音声認識モードにおける表示形式を示す図である。

【図 5】記録されたファイルの一覧表示形式を示す図である。

【図 6】テキストの表示形式を示す図である。

【図 7】画像の表示形式を示す図である。

【図 8】画像ファイルと、音声ファイル及びテキストファイルとが相互にリンクされた場合の双方のファイル内部の編成を示す図である。

【図 9】一覧表示において複数のファイルが選択された場合の表示形式を示す図である。

【図 10】グループファイルの内部データ編成である。＊

＊【図 11】グループに属する画像、音声、テキスト各ファイルの内部のデータ編成を示す図である。

【図 12】グループに属するファイルの一覧表示形式を示す図である。

【図 13】画像表示形式の上に音声認識制御パネルを重ねて表示したときの表示形式を示す図である。

【符号の説明】

1 撮影レンズ

6 撮像素子

10, 24 A/D変換器

12 バッファメモリ

13 信号処理制御用DSP

14 操作表示部

15 操作部

19 撮像信号処理部

20 音声入力回路

22 音声出力部

23 映像出力部

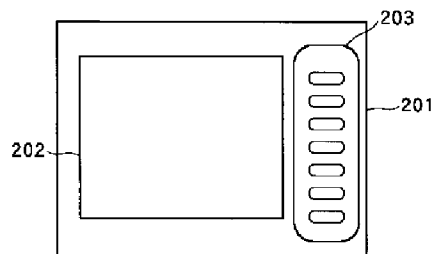
25, 26 D/A変換器

20 101 記録媒体

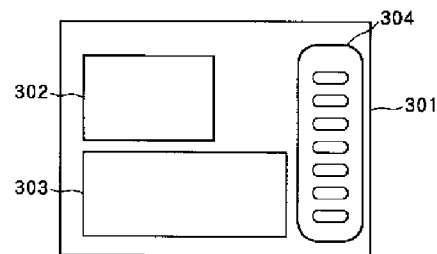
102 メモリバスコントローラ

104 記録媒体I/F

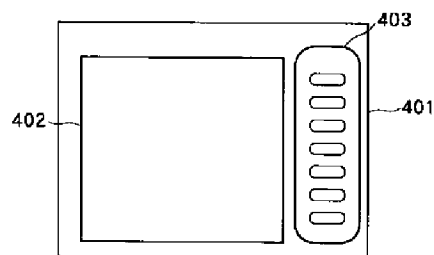
【図 2】



【図 3】



【図 4】

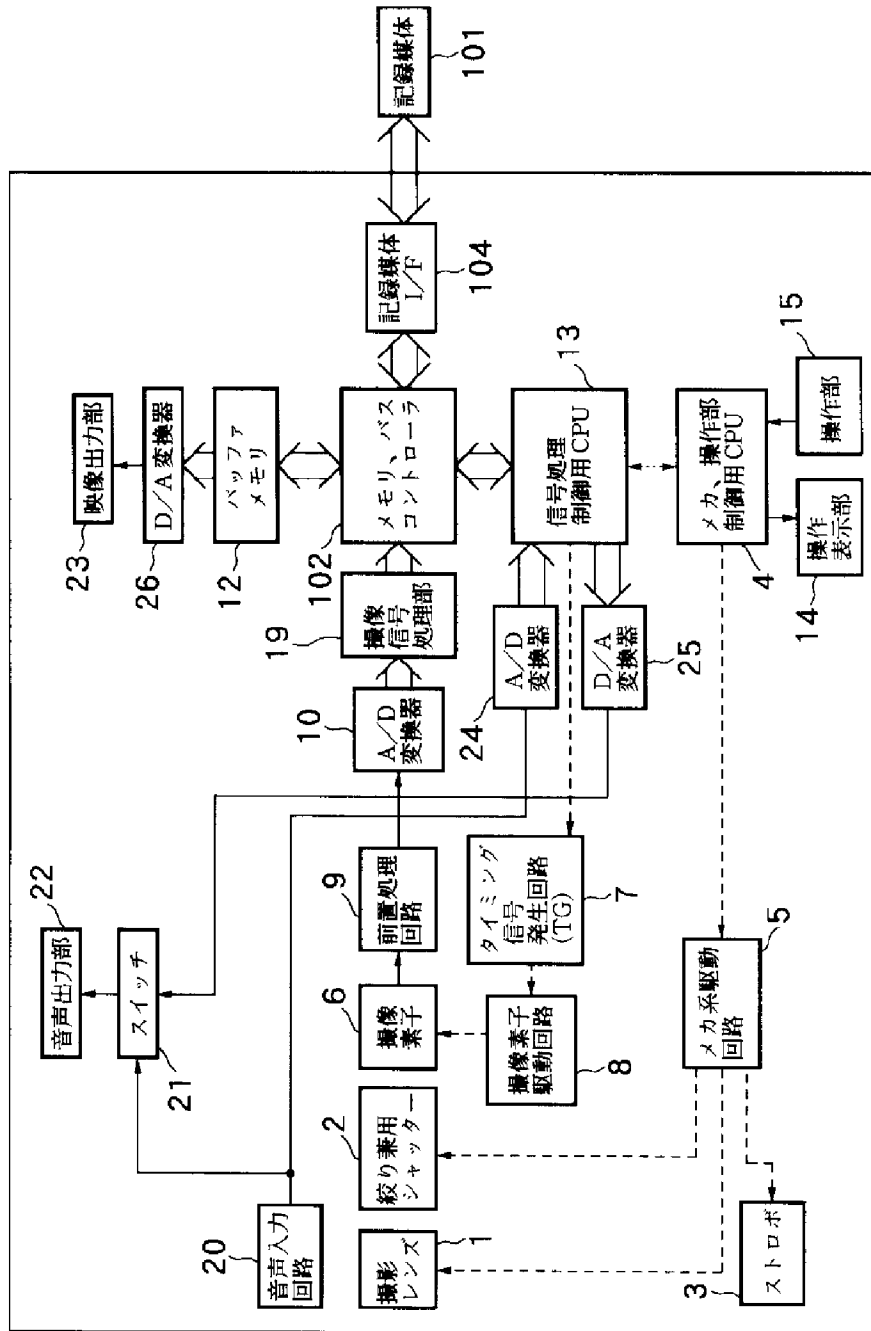


【図 11】

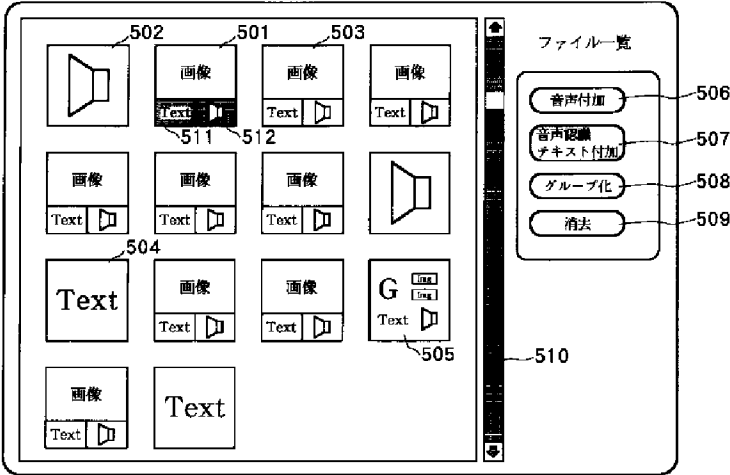
画像ファイル、
音声ファイル、
テキストファイル

リンクされたグループファイルの個数 p
グループファイル 1 の識別子
グループファイル 2 の識別子
グループファイル p の識別子

【図1】



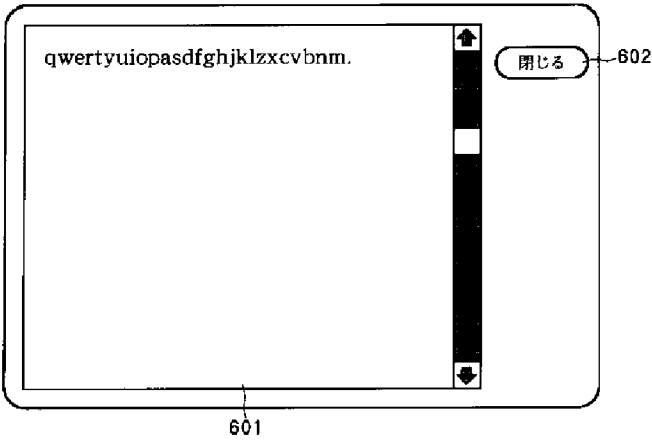
【図5】



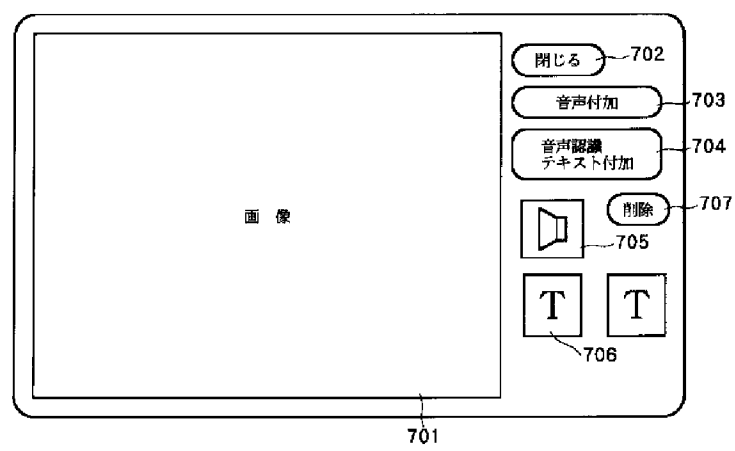
【図10】

グループファイル	
リンクされた画像ファイルの個数k	
画像ファイル1の識別子	
画像ファイル2の識別子	
画像ファイルkの識別子	
リンクされた音声ファイルの個数n	
音声ファイル1の識別子	
音声ファイル2の識別子	
音声ファイルnの識別子	
リンクされたテキストファイルの個数m	
テキストファイル1の識別子	
テキストファイル2の識別子	
テキストファイルmの識別子	

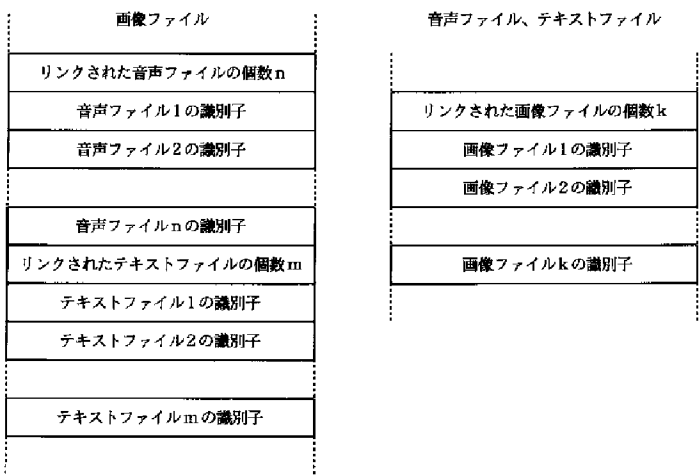
【図6】



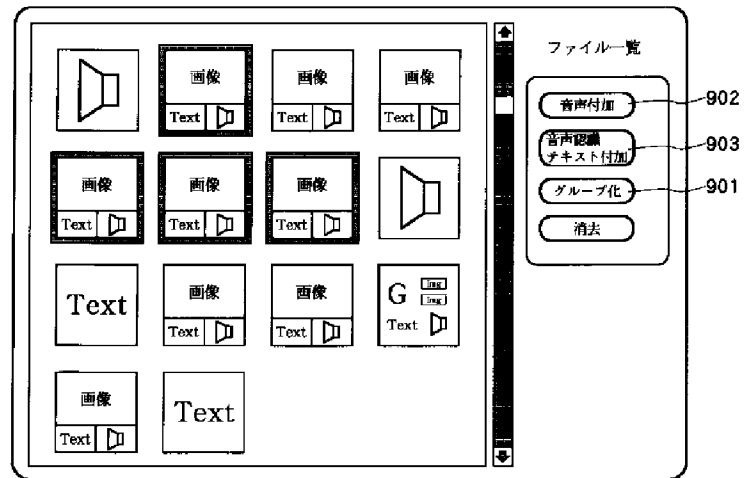
【図 7】



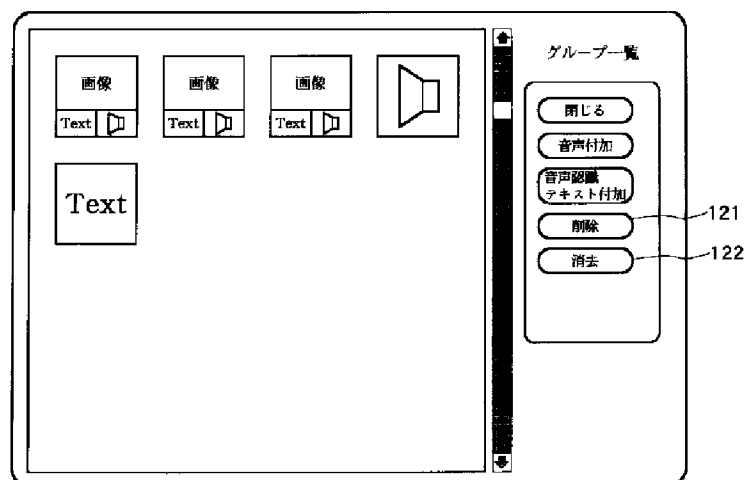
【図 8】



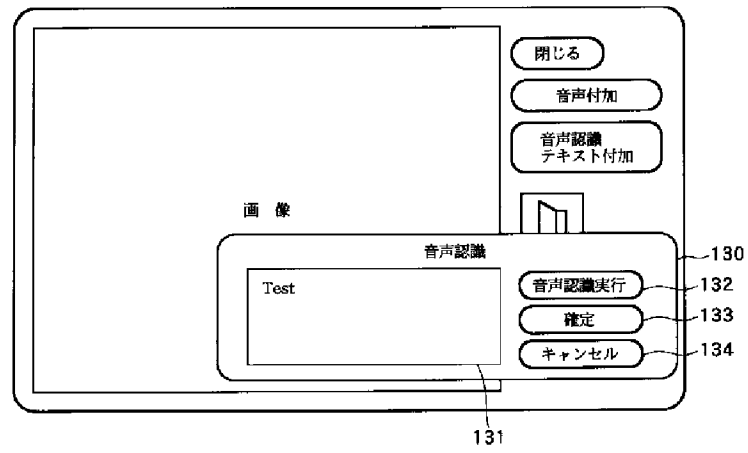
【図 9】



【図 12】



【図 13】



【公報種別】特許法第 17 条の 2 の規定による補正の掲載
【部門区分】第 7 部門第 3 区分
【発行日】平成 13 年 10 月 26 日 (2001. 10. 26)

【公開番号】特開平 7-184160
【公開日】平成 7 年 7 月 21 日 (1995. 7. 21)
【年通号数】公開特許公報 7-1842
【出願番号】特願平 5-328033
【国際特許分類第 7 版】

HD4N 5/91
5/781

【F I】

HD4N 5/91 R
5/781 C

【手続補正書】

【提出日】平成 12 年 12 月 25 日 (2000. 12. 25)

【手続補正 1】

【補正対象書類名】明細書

【補正対象項目名】発明の名称

【補正方法】変更

【補正内容】

【発明の名称】画像データ及び音声データを処理する装置、データ処理装置並びにデータ処理方法

【手続補正 2】

【補正対象書類名】明細書

【補正対象項目名】特許請求の範囲

【補正方法】変更

【補正内容】

【特許請求の範囲】

【請求項 1】 画像データと音声データとを関連づけて記録するための記録媒体と、
前記記録媒体に格納された音声データを指定する指定手段と、
前記指定された音声データに関連する画像データをファイル情報として検索する制御手段と、
前記検索された画像データを読み出し、再生する読出手段とを具備することを特徴とする画像データ及び音声データを処理する装置。

【請求項 2】 前記画像データから文字認識を行なう文字認識手段を更に具備し、該文字認識手段による認識結果をテキストファイルとして前記画像及び音声の各データとに関連づけて記録し、かつ検索できるように登録処理を行うことを特徴とする請求項 1 に記載の画像データ及び音声データを処理する装置。

【請求項 3】 前記音声データを文字として認識する音声認識処理手段を更に具備し、該音声認識処理手段による認識結果をテキストファイルとして前記画像及び音声の各データとに関連づけて記録し、かつ検索できるように

に登録処理を行うことを特徴とする請求項 1 に記載の画像データ及び音声データを処理する装置。

【請求項 4】 前記記録媒体に格納された前記各データをファイル情報として表示する表示手段を更に具備し、前記各ファイル間での相互関連情報の編集、及び該関連情報に基づく前記各データの検索結果を表示することを特徴とする請求項 1 に記載の画像データ及び音声データを処理する装置。

【請求項 5】 画像データに対応するアイコンと、前記画像データにリンクされた音声データに対応するアイコンとを近接して表示することにより、前記画像データと音声データのリンク関係を表示する表示手段と、
前記アイコンの一方を消去することによって、前記リンクを同時に消去する消去手段とを具備することを特徴とするデータ処理装置。

【請求項 6】 前記表示手段は、前記画像データあるいは音声データにリンクされたテキストデータに対応するアイコンを、前記画像データあるいは音声データのアイコンに近接して表示することを特徴とする請求項 5 に記載のデータ処理装置。

【請求項 7】 前記表示手段は、前記アイコンの表示されている同一画面内に、前記消去手段を動作させるための消去ボタンのアイコンを表示することを特徴とする請求項 6 に記載のデータ処理装置。

【請求項 8】 画像データと音声データとを関連づけて記録するための記録媒体に格納された音声データを指定する工程と、
前記指定された音声データに関連する画像データをファイル情報として検索する工程と、
前記検索された画像データを読み出し、再生する読出工程とを備えることを特徴とするデータ処理方法。

【請求項 9】 前記画像データから文字認識を行なう文字認識工程と、
前記文字認識工程における認識結果をテキストファイル

として前記画像及び音声の各データとに関連づけて記録し、かつ検索できる登録処理を行う工程とを更に備えることを特徴とする請求項 8 に記載のデータ処理方法。

【請求項 10】 前記音声データを文字として認識する音声認識処理工程と、前記音声認識処理工程における認識結果をテキストファイルとして前記画像及び音声の各データとに関連づけて記録し、かつ検索できるように登録処理を行う工程とを更に備えることを特徴とする請求項 8 に記載のデータ処理方法。

【請求項 11】 前記記録媒体に格納された前記各データをファイル情報として表示する表示工程を更に備え、前記各ファイル間での相互関連情報の編集、及び該関連情報に基づく前記各データの検索結果を表示することを特徴とする請求項 8 に記載のデータ処理方法。

【請求項 12】 画像データに対応するアイコンと、前記画像データにリンクされた音声データに対応するアイコンとを近接して表示することにより、前記画像データと音声データのリンク関係を表示する表示工程と、前記アイコンの一方を消去することによって、前記リンクを同時に消去する消去工程とを備えることを特徴とするデータ処理方法。

【請求項 13】 前記表示工程は、前記画像データあるいは音声データにリンクされたテキストデータに対応するアイコンを、前記画像データあるいは音声データのアイコンに近接して表示することを特徴とする請求項 12 に記載のデータ処理方法。

【請求項 14】 前記表示工程は、前記アイコンの表示されている同一画面内に、前記消去工程を実行させるための消去ボタンのアイコンを表示することを特徴とする請求項 12 に記載のデータ処理方法。

【手続補正 3】

【補正対象書類名】明細書

【補正対象項目名】0010

【補正方法】変更

【補正内容】

【0010】また、好ましくは、前記記録媒体に格納された前記各データをファイル情報として表示する表示手段を更に具備し、前記各ファイル間での相互関連情報の編集、及び該関連情報に基づく前記各データの検索結果を表示することを特徴としている。本発明のデータ処理装置は、画像データに対応するアイコンと、前記画像データにリンクされた音声データに対応するアイコンとを近接して表示することにより、前記画像データと音声データのリンク関係を表示する表示手段と、前記アイコンの一方を消去することによって、前記リンクを同時に消去する消去手段とを具備する。また、好ましくは、前記表示手段は、前記画像データあるいは音声データにリンクされたテキストデータに対応するアイコンを、前記画

像データあるいは音声データのアイコンに近接して表示する。また、好ましくは、前記表示手段は、前記アイコンの表示されている同一画面内に、前記消去手段を動作させるための消去ボタンのアイコンを表示する。本発明のデータ処理方法は、画像データと音声データとを関連づけて記録するための記録媒体に格納された音声データを指定する工程と、前記指定された音声データに関連する画像データをファイル情報として検索する工程と、前記検索された画像データを読み出し、再生する読出工程とを備える。また、好ましくは、前記画像データから文字認識を行なう文字認識工程と、前記文字認識工程における認識結果をテキストファイルとして前記画像及び音声の各データとに関連づけて記録し、かつ検索できる登録処理を行う工程とを更に備える。また、好ましくは、前記音声データを文字として認識する音声認識処理工程と、前記音声認識処理工程における認識結果をテキストファイルとして前記画像及び音声の各データとに関連づけて記録し、かつ検索できるように登録処理を行う工程とを更に備える。また、好ましくは、前記記録媒体に格納された前記各データをファイル情報として表示する表示工程を更に備え、前記各ファイル間での相互関連情報の編集、及び該関連情報に基づく前記各データの検索結果を表示する。本発明のデータ処理方法は、画像データに対応するアイコンと、前記画像データにリンクされた音声データに対応するアイコンとを近接して表示することにより、前記画像データと音声データのリンク関係を表示する表示工程と、前記アイコンの一方を消去することによって、前記リンクを同時に消去する消去工程とを備える。また、好ましくは、前記表示工程は、前記画像データあるいは音声データにリンクされたテキストデータに対応するアイコンを、前記画像データあるいは音声データのアイコンに近接して表示する。また、好ましくは、前記表示工程は、前記アイコンの表示されている同一画面内に、前記消去工程を実行させるための消去ボタンのアイコンを表示する。

【手続補正 4】

【補正対象書類名】明細書

【補正対象項目名】0037

【補正方法】変更

【補正内容】

【0037】＜テキストの再生＞

更に、テキストを表示したい場合、ユーザは操作部のポインティングデバイスを用いてテキストアイコンを選択して起動する。図 6 は、テキストデータの表示形式を示す図である。信号処理制御用 DSP 13 は記録媒体 I/F 104、メモリバスコントローラ 102 を制御してテキストデータを読み出しキャラクタビットパターン展開して例えば図 6 に示すようなテキストデータを画像表示用バッファメモリに転送し画像表示部に表示する。